

Multisystemic Treatment: A Meta-Analysis of Outcome Studies

Nicola M. Curtis and Kevin R. Ronan
Massey University

Charles M. Borduin
University of Missouri—Columbia

Multisystemic treatment (MST) is a family- and home-based therapeutic approach that has been found to be effective in treating antisocial youths and that has recently been applied to youths with serious emotional disturbances. In light of the increasing dissemination of MST, this review examines the effectiveness of MST by quantifying and summarizing the magnitude of effects (treatment outcomes) across all eligible MST outcome studies. Included in a meta-analysis were 7 primary outcome studies and 4 secondary studies involving a total of 708 participants. Results indicated that across different presenting problems and samples, the average effect of MST was $d = .55$; following treatment, youths and their families treated with MST were functioning better than 70% of youths and families treated alternatively. Results also showed that the average effect of MST was larger in studies involving graduate student therapists (i.e., efficacy studies; $d = .81$) than in studies with therapists from the community (i.e., effectiveness studies; $d = .26$). In addition, MST demonstrated larger effects on measures of family relations than on measures of individual adjustment or peer relations.

keywords: antisocial behavior, community-based treatment, juvenile offenders, multisystemic treatment, treatment adherence, youths

Antisocial behavior in youths represents a complex and pervasive clinical problem, with large numbers of antisocial youth coming to the attention of mental health, social welfare, and youth justice systems throughout the western world each year (Kazdin, 1987; Rutter, Giller, & Hagell, 1998). Recent prevalence statistics from the United States, the United Kingdom, and New Zealand indicate that antisocial behaviors manifest in up to 15% of young people (Fergusson, Horwood, & Lynskey, 1997; Office of Juvenile Justice and Delinquency Prevention, 1999).

Treatment approaches have typically focused on different aspects of the range of dysfunction (e.g., individual or family factors) found in antisocial youths. Despite a large number of available treatments, few have demonstrated sustained effectiveness in the amelioration of serious and pervasive antisocial behavior (Kazdin, 2000). However, leading reviewers (e.g., Burns, Hoagwood, & Mrazek, 1999; Elliott, 1998; Farrington & Welsh, 1999; Kazdin & Weisz, 1998; U.S. Public Health Service, 2001) of empirically supported child and adolescent treatments have noted that multisystemic treatment (MST) was effective across various replications, problems, therapists, and settings.

MST is a treatment model that emphasizes recognized

risk factors associated with antisocial behavior. MST is underpinned by both Bronfenbrenner's (1979) theory of social ecology and the empirically validated determinants of antisocial behavior in youth (Borduin & Schaeffer, 1998; Lahey, Moffitt, & Caspi, 2003; Loeber, Farrington, & Waschbusch, 1998). MST interventions target the individual, family, peer, school, and community factors identified as contributing to and maintaining problematic behavior. Interventions aim to empower parents to facilitate pragmatic changes in the youth's and the family's natural environments (home and other community-based settings).

Ongoing evaluation of outcomes has been an important feature in the development of MST. Serious juvenile offenders have been the primary focus of MST outcome studies conducted to date (Borduin et al., 1995; Henggeler, Melton, & Smith, 1992; Henggeler et al., 1997). Other studies have used MST in comorbid populations (i.e., delinquency comorbid with substance abuse and/or severe emotional disturbance) (Henggeler, Pickrel, & Brondino, 1999; Henggeler, Rowland, et al., 1999). A quality assurance system has been implemented in more recent outcome studies of MST in an attempt to ensure that treatment fidelity is maintained in the absence of the treatment developers and across a range of applied settings.

Despite the encouraging research outcomes that have been reported for MST in recent narrative reviews (Borduin, 1999; Borduin, Schaeffer, & Ronis, 2003; Brown, Borduin, & Henggeler, 2001), no systematic quantitative review of this body of research has been conducted. Given the increasing dissemination of MST in the United States and abroad,

Nicola M. Curtis and Kevin R. Ronan, School of Psychology, Massey University, Palmerston North, New Zealand; Charles M. Borduin, Department of Psychological Sciences, University of Missouri—Columbia.

Correspondence concerning this article should be addressed to Nicola M. Curtis or Kevin R. Ronan, School of Psychology, Massey University, Private Bag 11-222, Palmerston North, New Zealand. E-mail: nmcurtis@xtra.co.nz or k.r.ronan@massey.ac.nz

it would be useful to provide quantitative summary data of overall treatment effectiveness¹ in the form of a meta-analysis. The present review examines the effectiveness of MST by quantifying and summarizing the magnitude of effects (treatment outcomes) across all eligible MST outcome studies. Categorical variables that may account for variation in treatment outcomes are also examined.

Literature Review Procedures

Literature Search

The time frame of the literature search spans the time period from 1986 (when controlled outcome research on MST began) through 2003. Ninety-one articles were identified from studies listed in the Psychological Literature and Educational Resources Information Center databases using the keywords *multisystemic therapy* and *multisystemic treatment* crossed with *treatment outcomes*, *juvenile delinquency*, *antisocial behavior*, and *family relations*. In addition, the recent tables of contents of journals² most likely to publish studies on MST were manually searched.

Selection Criteria

Inclusion of studies in the meta-analysis required (a) identification of the treatment approach as MST, including documented adherence to the MST treatment principles (Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998); (b) random assignment of participants to MST and one or more control groups; (c) a clinical sample in which youths or their parents/caregivers manifested antisocial behavior (defined as social rule violations, acts against others, or both) and/or psychiatric symptoms; (d) use of both pretreatment and posttreatment assessment measures and/or follow-up assessment measures; and (d) use of test statistics suitable for meta-analysis (means, standard deviations, and/or Fisher's *F* ratios).

Search Outcome

Overall, seven primary outcome studies containing a total of 708 participants and 35 MST therapists met inclusionary criteria and were included in this meta-analysis. These studies were all published in peer-reviewed journals between 1987 and 2002. One other primary study was not included in the meta-analysis because insufficient test statistics were reported (Henggeler et al., 1986). Four secondary studies (i.e., studies reporting secondary analyses of data from primary outcome studies) were also included (Brown, Henggeler, Schoenwald, Brondino, & Pickrel, 1999; Henggeler et al., 1991; Henggeler, Clingempeel, Brondino, & Pickrel, 2002; Schoenwald, Ward, Henggeler, & Rowland, 2000). Three other secondary studies were not included because insufficient test statistics were reported (i.e., Henggeler, Melton, Smith, Schoenwald, & Hanley, 1993; Schoenwald, Ward, Henggeler, Pickrel, & Patel, 1996) or data were reported from a small subsample of a later pri-

mary study (i.e., Scherer, Brondino, Henggeler, Melton, & Hanley, 1994).

Statistical Procedures

Effect sizes. To derive information about the magnitude of the differences between treatment groups, we expressed comparisons in terms of a standardized measure of effect size, the *d* index (Cohen, 1977). The *d* index is defined for present purposes as the difference between the mean change scores of two groups divided by the average or common standard deviation of the groups. This calculation results in a measure of the degree to which two groups differ in standard deviation units.

Effect sizes were calculated from three primary studies that used an alternative treatment as the control condition and from four primary studies that included a usual services control group. Effect sizes from secondary studies ($n = 4$) were only included when the outcomes were not reported in the related primary study. In studies in which means and standard deviations were not reported ($n = 2$), effect sizes (*d* indexes) were estimated by converting reported *F* values (Cooper, 1998).

Correction for bias. Effect sizes based on small samples tend to inflate the population values that they estimate and must be reduced (Lipsey, 1992). To account for small sample bias ($N < 30$), the effect size of one study (Borduin, Henggeler, Blaske & Stein, 1990) was adjusted using the correction procedure recommended by Hedges (1991). The specific weighting coefficient used for effect size adjustments in this study was $1 - (3/4n_t + 4n_c - 9)$, where n_t is the sample size for the treatment group and n_c is the sample size for the control group (Hedges & Olkin, 1985). No other attempts were made to adjust for sample size at the time of measurement (as recommended by Lipsey, 1992).

Confidence levels. Ninety-five percent confidence intervals (CIs) are reported for the overall effect size estimate. The standard error was estimated by applying the following formula to each study's effect size (Hedges & Olkin, 1985):

¹ Although all of the MST outcome studies were conducted in community settings with real-world clinical samples, three of the seven studies (Borduin, Henggeler, Blaske, & Stein, 1990; Borduin et al., 1995; Brunk, Henggeler, & Whelan, 1987) involved graduate students as therapists and thus should probably be classified as "efficacy" (rather than "effectiveness") studies (see Bickman & Noser, 1999). Even so, for economy of expression, we use the term *effectiveness* throughout the article when referring to the overall outcomes of MST.

² Recent volumes of the following journals were hand searched: *Journal of the American Academy of Child and Adolescent Psychiatry*, *Journal of Consulting and Clinical Psychology*, *Journal of Counseling Psychology*, *Psychological Bulletin*, *Journal of Child and Family Studies*, *American Journal of Psychiatry*, *Journal of Emotional and Behavioral Disorders*, *Mental Health Services Research*, *International Journal of Offender Therapy*, and *Comparative Criminology*.

$$CI = [d - 1.96(SE)] \text{ to } [d + 1.96(SE)],$$

where $SE = 1/\sqrt{\sum w_i}$

$$w_i = \left[\frac{1}{v_i} \right] \text{ and } v_i = \left[\frac{n_i \pm n_c}{n_i n_c} + \frac{d_2}{2(n_i + n_c)} \right].$$

Statistical power. The power of a statistical test is defined as the probability that it will yield a true effect that is statistically significant (i.e., reducing the likelihood of making a Type II error; Cohen, 1988). In a meta-analysis, a power survey estimates the proportion of studies that should yield a statistically significant effect (Borenstein, Rothstein, & Cohen, 1997). Statistical power was calculated based on sample sizes, Cohen's (1988) power tables, and effect size conventions ($d = .20$ for small effects, $d = .50$ for medium effects, and $d = .80$ for large effects).

Homogeneity of effect sizes. The Q_i statistic was computed to test for homogeneity among primary study outcomes. This statistic evaluates whether all studies have the same population effect size (i.e., whether the variation in effect sizes is no greater than would be expected due to errors in sampling or measurement; Hedges & Olkin, 1985). Q_i is distributed as a chi-square variable with $K - 1$ degrees of freedom, where K equals the number of effect sizes. If the Q_i is not significant, the reviewer can assume that the effect sizes reported for the group of studies are homogeneous. If the Q_i is significant, the effect sizes are considered to be heterogeneous, and the reviewer should try to determine which studies (or effect sizes) might be included in further subsets of the studies. In other words, the studies can be partitioned into groups of effect sizes according to the theoretical or practical importance of the grouping variable.

Assuming that the Q_i is significant and that studies can be partitioned into meaningful groups, two other tests are then used to evaluate possible differences between the groups. First, the Q_b statistic (which has an approximate chi-square distribution with $p - 1$ degrees of freedom, where p equals the number of categories or groups) is used to test whether the average effect sizes from the groupings are homogeneous (Cooper, 1998). If the Q_b is not significant, then the average d indexes are considered homogeneous and the grouping factor does not explain variance in effects beyond that associated with sampling error. If Q_b exceeds the critical value (i.e., is significant), then the grouping factor is a significant contributor to variance in effect sizes. However, Q_b can only be interpreted correctly in conjunction with a second statistic, Q_w . The Q_w statistic (distributed as a chi-square with $K - p$ degrees of freedom) is used to provide an estimate of within-class homogeneity. As recommended by Lipsey and Wilson (2001), a correctly specified grouping variable (i.e., categorical moderator) that accounts for the heterogeneity among effect sizes across studies is achieved when the value of Q_b is significant (i.e., mean d differs between and/or among groups) and the value of Q_w is not significant (i.e., no heterogeneity remains unmodeled given

the moderator and the conditional variances/weights that quantify random subject sampling).

Results

Characteristics of Participants

All studies were conducted in the United States, and the primary studies were funded through local, state, and/or federal mental health agencies ($n = 6$) or by a research center of excellence ($n = 1$). Study sample sizes ranged from 16 to 176, with a median of 116. The youths ranged in age from 8.3 to 17.6 years ($Mdn = 14.8$), 70% were male, and 81% lived with at least one biological parent. Fifty-four percent ($n = 380$) of the youths were African American, 45% ($n = 319$) were Caucasian, 0.7% ($n = 5$) were Hispanic American, and 0.5% ($n = 4$) were Asian American. Fifty-nine percent ($n = 415$) of the youths were classified as chronic, at-risk, and/or violent juvenile offenders; 17% ($n = 118$) were classified as substance abusers; 16% ($n = 116$) required emergency psychiatric hospitalization (presenting problems included suicidal ideation, homicidal ideation, and psychosis); 6% ($n = 43$) were classified as abused (including physical abuse and psychological trauma) and/or neglected (including abandonment/lack of supervision and inadequate care); and 2% ($n = 16$) were classified as sexual offenders. Eighty-four percent ($n = 593$) of the youths had been arrested previously. Insufficient information was provided in the primary studies to derive an overall socioeconomic score according to Hollingshead's (1975) criteria. However, on the basis of information regarding parental education, employment status, single-parent status, and median income, it appears that the samples in most studies were drawn from disadvantaged populations.

Characteristics of Treatments and Therapists in Primary MST Studies

MST was compared with a range of usual services in four studies and with other treatment programs in the remaining three studies (see Table 1). Usual services were provided through (a) juvenile justice agencies, (b) a community mental health center, (c) an outpatient substance abuse treatment program, and (d) an inpatient psychiatric hospital. Youths assigned to juvenile justice agencies were monitored for school attendance and were seen weekly, fortnightly, or monthly by probation officers for up to 6 months; these youths were also referred to other social service agencies (i.e., substance abuse treatment agencies, community mental health agencies) as necessary. Youths in the community mental health comparison group received family or individual counseling, social skills training, and/or vocational training. The outpatient substance abuse service offered adolescent group therapy. Youths in the inpatient hospitalization group were provided with crisis stabilization, psychiatric evaluation, and intensive individualized care. Across studies, youths in usual services conditions received

Table 1
Clinical Population, Comparison Condition, and Mean Effect Size for Multisystemic Treatment Outcome Studies

Study	<i>N</i>	Population	Comparison condition	<i>d</i>	<i>SD</i>
1. Brunk et al. (1987)	43	Abusing/neglectful parents	Parent training	1.32	0.65
2. Borduin et al. (1990)	16	Juvenile sexual offenders	Individual therapy	1.08	0.23
3. Henggeler et al. (1991)		Same sample as Studies 4 and 5		0.64	0.33
4. Henggeler et al. (1992)	84	Violent and chronic juvenile offenders	Individual therapy	0.37	0.13
5. Borduin et al. (1995)	176	Violent and chronic juvenile offenders	Individual therapy	0.66	0.43
6. Henggeler et al. (1997)	155	Violent and chronic juvenile offenders	Usual services	0.27	0.25
7. Henggeler, Pickrel, et al. (1999)	118	Substance abusing juvenile offenders	Usual services	0.25	0.08
8. Henggeler, Rowland, et al. (1999)	116	Psychiatrically disturbed adolescents	Usual services	0.19	0.92
9. Brown et al. (1999)		Same sample as Study 7		0.60	0.36
10. Schoenwald, Ward, et al. (2000)		Same sample as Study 9		0.52	0.22
11. Henggeler, Clingempeel, et al. (2002)		Same sample as Study 7		0.15	0.02

Note. The *d* for each primary outcome study is an average effect across multiple outcome measures. The *d* for each secondary analysis of data is an average effect across multiple outcome measures not reported in primary outcome studies.

an average of 20 more hours of services than did youths in MST.

The three comparison treatment programs included parent training ($n = 1$ study) and individual therapy ($n = 2$ studies). Parent training consisted of weekly group sessions in which caregivers received instruction on human development, behavioral management techniques, and positive parent-child interactions. Individual therapy for the youths included an eclectic blend of psychodynamic, client-centered, and behavioral approaches in which therapists focused on personal, family, and academic issues and provided encouragement for behavior change. Across studies, youths in individual therapy conditions received an average of 6.3 hr more treatment than did youths in MST. Treatment completion rates ranged from 76% to 100% for MST and from 56% to 100% for other treatment conditions. The average treatment completion rate over primary and secondary studies was 86% for MST and 78% for other treatments.

MST was most often conducted in family homes and other community settings (e.g., schools). MST was of brief duration, averaging approximately 40 hr over 15 weeks for up to 24 weeks. Ninety-seven percent of the MST therapists were either current graduate students (with a bachelor's or master's degree) or had earned a terminal master's degree. MST therapists had 1 to 15 years of clinical experience in social work, pastoral counseling, psychology, or other related mental health fields. MST therapists received 3 to 6 days of intensive didactic and experiential training and attended additional quarterly booster training sessions. Among MST therapists, 43% ($n = 15$) were male, 37% ($n = 13$) were African American, and 63% ($n = 22$) were Caucasian. MST therapists received an average of 1.7 hr of supervision each week.

Outcome Measures

A multiagent, multimethod assessment approach was taken in six of seven studies, with a total of 23 different outcome measures being used ($M = 6.4$ per study; see Table 2). The one study that did not include multiple assessment

measures (Borduin et al., 1990) determined treatment outcomes from re-arrest data obtained from juvenile court, adult court, and state police records.

Ultimate goals, which are common to all treatments of juvenile offenders, were assessed across studies and included changes in (a) the rate, frequency, and seriousness of adolescent criminal activity; (b) days incarcerated; (c) days absent from school; (d) alcohol and marijuana use; and (e) days in mandated out-of-home placements (including hospitalization). Ultimate outcomes were typically assessed at posttreatment (i.e., treatment completion) and follow-up assessments, the latter of which were conducted from 12 weeks to 4 years following treatment completion.

Treatment Adherence

Therapist adherence to the MST treatment protocol was assessed with the 26-item MST Adherence Measure (Henggeler & Borduin, 1992) in three of the studies (Henggeler, Pickrel, et al., 1999; Henggeler et al., 1997; Henggeler, Rowland, et al., 1999). Items on the measure assess six factors that reflect (a) therapist adherence to the MST treatment principles, (b) the degree to which therapy sessions were nonproductive, (c) problem-solving efforts of the therapist and family, (d) therapist attempts to change family interactions, (e) lack of therapeutic direction in sessions, and (f) the degree of family therapist consensus. The measure was administered to families and therapists following randomly selected sessions during the 4th and 8th weeks of treatment.

Magnitude of Effects

Effect sizes ranged in magnitude from -0.02 to 5.79 . As recommended by Cooper (1998), both significant and non-significant *d* index values were included in the analyses in an attempt to minimize bias and maximize confidence in any conclusions that were drawn. There were a total of 101 *d* index values, and 6 of the 7 studies had multiple indexes.

Table 2
Mean Effect Size for Domain and Source of Outcome Measure

Outcome	Effect size (<i>d</i>)		Relevant studies
	Domain	Subdomain (variable)	
Individual	.28		2, 4, 6, 7, 8, 9, 10
Youth symptoms		.43 ↓	2, 4, 9, 10
Parent symptoms		.33 ↓	2, 9, 10
Youth behavior problems		.34 ↓	2, 7, 9
Hospitalization		.52 ↓	11
Family	.57		2, 4, 7, 9, 10
Self-reported family relations			
Adaptability/cohesion		.31 ↑	2, 7, 9, 10
Parental monitoring		.60 ↑	2
Stress		1.01 ↓	4
Observed family interactions	.76		2, 4
Conflict-hostility		.62 ↓	2, 4
Overall family supportiveness		.84 ↑	2, 4
Parental effectiveness		.94 ↑	2, 4
Verbal activity		.22 ↑	2, 4
Youth noncompliance		.92 ↓	2, 4
Peer relations	.11		2, 7, 9, 10
Aggression		.02 ↓	2, 7, 9
Bonding		.08 ↑	2, 7, 9
Social maturity		.07 ↑	2, 9
Social competence		.28 ↑	7, 10
Association with deviant peers		.31 ↓	9, 10
School attendance	.54		3, 10
Ultimate outcomes (criminal activity)	.50		1, 2, 6, 7, 8, 9, 10
Number of arrests for all crimes		.55 ↓	1, 2, 7, 8, 9, 10
Number of arrests for substance abuse crimes		.29 ↓	6, 7, 10
Seriousness of arrests		1.01 ↓	2
Days incarcerated		.55 ↓	7, 9
Self-reported delinquency		.07 ↓	8, 9
Self-reported drug use		.64 ↓	5

Note. Upward and downward arrows indicate a respective increase or decrease in the associated domain. 1 = Borduin et al. (1990); 2 = Borduin et al. (1995); 3 = Brown et al. (1999); 4 = Brunk et al. (1987); 5 = Henggeler et al. (1991); 6 = Henggeler, Clingempeel, et al. (2002); 7 = Henggeler et al. (1992); 8 = Henggeler, Pickrel et al. (1999); 9 = Henggeler et al. (1997); 10 = Henggeler, Rowland et al. (1999); 11 = Schoenwald, Ward, et al. (2000).

These *d* index values were averaged to yield one *d* index per study.

The average effect of MST across the seven primary and four secondary outcome studies was $d = .55$ ($Mdn = .52$). The significance of this effect size was tested by computing a 95% confidence interval, which ranged from $d = .40$ to $d = .70$. As the lower limit of the 95% confidence interval is well above zero, the mean *d* index value is assumed to be significantly different from zero (Shadish & Haddock, 1994). Of note, for 6 of the 101 *d* indexes (6%), improvements in the control group exceeded those in the MST group.

U_3 is a measure of distribution overlap that provides another method of interpreting the *d* index (Cohen, 1988). U_3 tells the percentage of people in the lower mean group who are surpassed by the average person in the higher mean group. In the present case, a *d* of .55 equates to a U_3 of .70,

meaning that the average participant in the MST condition surpassed 70% of the control condition participants on the measures of instrumental and ultimate outcomes.

Power Analysis

The *d* of .55 calculated in this study can be categorized as a moderate effect (see Cohen, 1988). On the basis of the overall comparison between MST ($n = 361$) and control ($n = 347$) groups, the power to detect a moderate effect size was .57 ($Mdn = .73$). Thus, on average, investigators had a 57% chance of detecting an effect size in the moderate range. In terms of detecting a moderate effect, power failed to reach the 80% criterion for design sensitivity (Cohen, 1988). Accordingly, as there is an increased likelihood of making a Type II error (i.e., rejecting the null hypothesis

when it is true), results of this analysis should be interpreted with some caution.

Relationship Between Effect Size and Treatment Domain

The effect sizes achieved in the different instrumental domains (individual, family, peer) that are targeted in the MST model were examined. A comparison of average effect sizes on measures of individual adjustment ($d = .28$) versus measures of peer relations ($d = .11$) did not reveal a significant difference ($Z = -0.92, ns$). However, measures of family relations demonstrated a larger average effect size ($d = .57$) than did measures of individual adjustment ($d = .28$) or measures of peer relations ($d = .11$; $Z_s > -2.13, ps < .03$). Table 2 summarizes the effect sizes for various domains and subdomains.

Categorical Moderator Analyses

The homogeneity analysis for the effect sizes in the present review revealed there was more variability in the combined d indexes than would be expected due to sampling error or other sources of expected error, $Q_b(df = 1) = 11.73, p < .05$. Potential moderator variables were then tested to identify sources of heterogeneity among studies.

Studies were initially grouped into mutually exclusive categories on the basis of the target population (i.e., violent and chronic juvenile offenders vs. all other youth populations, including psychiatrically disturbed youths, substance-abusing juvenile offenders, juvenile sexual offenders, and abused or neglected youths). Three of the seven studies (i.e., Borduin et al., 1995; Henggeler et al., 1992, 1997) used MST with populations of violent and chronic juvenile offenders. The average effect size achieved in these studies ($d = .44, CI = \pm .19$) was compared with the average effect size achieved by the remaining studies ($d = .38, CI = \pm .27$). The analyses of differences in effect sizes across studies, $Q_b(df = 1) = .11, ns$, and within studies, $Q_w(df = 5) = 11.62, p < .05$, indicated that the type of target population did not moderate treatment effects.

Given that target population did not account for the heterogeneity between studies, we then examined whether differences in study conditions (i.e., efficacy vs. effectiveness conditions) might moderate treatment effects. Although all of the MST outcome studies were conducted in community settings, three of the seven studies (i.e., Borduin et al., 1990, 1995; Brunk, Henggeler, & Whelan, 1987) used closely supervised graduate students as therapists. Such conditions are more characteristic of "efficacy" studies than "effectiveness" studies (see Bickman & Noser, 1999). That is, more control was exercised over the treatment conditions than in the remaining four studies (Henggeler et al., 1992; Henggeler, Pickrel, & Brondino, 1999; Henggeler et al., 1997; Henggeler, Rowland et al., 1999). The average effect size achieved in more controlled studies with graduate student therapists ($d = .81, CI = \pm .33$) was compared with the average effect size achieved in community-based studies

($d = .26, CI = \pm .06$). The computed values of the tests of differences in effect sizes across studies, $Q_b(df = 1) = 8.74, p < .05$, and within studies, $Q_w(df = 5) = 2.99, ns$, indicated that the effect sizes were not homogeneous across groups. Thus, the study conditions variable may account for the heterogeneity among studies.³

Discussion

The primary objective of this review was to establish an estimate of the overall effectiveness of MST in treating antisocial and associated behaviors in a range of youth and family populations. Across both instrumental and ultimate outcome measures, youths and their families treated with MST were functioning better and offending less than 70% of their counterparts who received alternative treatment or services. Specifically, MST was found to be relatively effective in reducing emotional and behavioral problems in individual family members, in improving parent-youth and overall family relations, in decreasing youth aggression toward peers and involvement with deviant peers, and in reducing youth criminality. Follow-up data suggest that treatment effects were sustained for up to 4 years.

This review found that MST demonstrated larger effects on measures of family relations than on measures of individual adjustment or peer relations. This finding is consistent with the emphasis that MST places on family interventions (Henggeler & Borduin, 1990) and with previous studies of change processes in MST showing that improvements in family relations predicted decreases in individual problems (i.e., symptoms, delinquent behavior) and in delinquent peer affiliation (Huey, Henggeler, Brondino, & Pickrel, 2000; Mann, Borduin, Henggeler, & Blaske, 1990). The larger observed effect of MST on family relations measures than on other measures was also likely due to the fact that the comparison conditions (i.e., usual services or an alternate treatment) typically focused on the individual adolescent and not on the family. Indeed, two of the studies included in this review (Borduin et al., 1995; Henggeler et al., 1992) showed that, in contrast to MST, the individually focused comparison conditions led to a deterioration in family relations over the course of treatment. This deterioration in family relations has also been observed in other studies of individually focused child and adolescent treatments (e.g., Szapocznik et al., 1989) and is consistent with

³ Another key difference between the "efficacy" and "effectiveness" studies pertained to the nature of the control groups that were used. More specifically, all of the efficacy studies included an alternative treatment control group (either individual therapy or parent training), whereas all of the effectiveness studies included a usual services control group. Inasmuch as study condition (efficacy vs. effectiveness) was confounded with the type of control group (alternative treatment vs. usual services) in these studies, it is not possible to ascertain that study condition per se accounted for the heterogeneity in effect sizes. It is possible that moderator effects can be attributed to both the type of study condition and the type of comparison group.

the systemic perspective that child misbehavior often serves a functional purpose (e.g., by uniting parents who are otherwise in conflict) in the family (Hoffman, 1981; Minuchin, 1985). From this perspective, treatments that focus primarily on improving the individual child's behavior may destabilize the family system by removing the child from his or her central position in family (or marital) conflicts.

The results of this review also indicate that treatment effect sizes in MST outcome studies were not moderated by the type of target population (broadly defined as violent and chronic juvenile offenders vs. other populations of youths) that participated in MST. MST was originally developed and validated with populations of serious and violent juvenile offenders and only more recently has been extended to populations of youths with substance abuse problems or serious emotional disturbances. Although the results of this review suggest that MST appears to be a promising approach for populations other than violent and chronic juvenile offenders, additional studies evaluating the effectiveness of MST with these other populations will be needed before more definitive conclusions can be drawn and before dissemination efforts would be justified.

Another issue highlighted by this review pertains to the dissemination of efficacious treatments to community settings. The results indicate that treatment effects in MST outcome studies might have been moderated by differences in study conditions (i.e., efficacy vs. effectiveness conditions). It is entirely possible that the involvement of the MST developers as clinical supervisors in the efficacy studies contributed to the higher effect sizes that were observed in those studies ($d = .81$) versus the effectiveness studies ($d = .26$). In fact, Schoenwald, Henggeler, Brondino, and Rowland (2000) have noted that ongoing quality assurance procedures (e.g., therapist and supervisor adherence protocols) are indispensable when disseminating MST to community settings. It is also important to note that most of the recent MST clinical trials have emphasized ongoing evaluation of various components of the MST quality assurance process. In particular, several studies have examined determinants (e.g., supervision by MST-trained supervisors, organizational support for the MST model) of therapist fidelity to the MST model (Henggeler et al., 1997; Henggeler, Schoenwald, Liao, Letourneau, & Edwards, 2002; Huey et al., 2000; Schoenwald & Hoagwood, 2001). It remains to be seen whether MST effectiveness studies will be able to demonstrate results that are comparable to those that have been obtained in MST efficacy studies.

In addition to improving fidelity to the MST model in community settings, researchers of MST effectiveness should broaden their assessment of instrumental outcomes in each of the systems pertinent to the goals of MST. For example, although a common goal of MST is to increase youth involvement with prosocial peers, the assessment of change in peer affiliations has been limited to measures of association with deviant peers. Measures that directly assess involvement with prosocial peers would be informative in future studies. Similarly, assessment of relevant cognitive processes such as attributional style or bias would be a

valuable addition to the assessment of individual behavior problems and psychiatric symptoms. Broader assessment of other areas that are frequent targets of MST interventions might include measures of performance in school (e.g., grades, achievement levels) and participation in extracurricular activities (e.g., sports teams, church groups, recreation center activities).

Research is also needed to evaluate the processes that MST incorporates to facilitate changes in youths and their families. Although investigators have begun to identify important moderators (e.g., treatment fidelity; Henggeler et al., 1997) and mediators (e.g., improved peer relations; Huey et al., 2000; family engagement; Schaeffer & Borduin, 2003) of MST outcomes, greater understanding of specific mechanisms of change is required to enable community-based agencies to apply targeted and cost-effective treatment approaches such as MST.

Findings of the present review must be interpreted in the context of the following methodological limitations. First, given the relatively small number of outcome studies that were available for inclusion in the review, the conclusions should be considered tentative. Nevertheless, a total of 708 participants helped to offset the limited number of separate investigations. Second, some of the subcategories of effects were based on a small number (i.e., subgroup) of studies, thus limiting the generalizability of the findings. Third, the confounding of study condition (efficacy vs. effectiveness) with type of control group (alternative treatment vs. usual services) in the moderator analysis clouds the interpretation of the results. This interpretive ambiguity cannot be resolved in the present study. Finally, it is prudent to acknowledge the potential for bias in the current review due to the pooling of sample studies (Hedges & Olkin, 1985). However, it is also the case that a more conservative approach to pooling effect sizes strengthens confidence in the results.

In conclusion, as an empirically established treatment for violent and chronic juvenile offenders, MST appears to be worthy of wider implementation and continued evaluation. The overarching objective of MST (i.e., empowering parents to facilitate pragmatic changes in the youth's and the family's natural environments) appears to be robust with this population. More empirical support is required before MST can be considered an effective treatment of substance abuse in adolescents or an effective community-based alternative to the hospitalization of youths presenting psychiatric emergencies.

References

- References marked with one or two asterisks indicate primary studies or secondary analysis studies, respectively, in the meta-analysis.
- Bickman, L., & Noser, K. (1999). Meeting the challenges in the delivery of child and adolescent mental health services in the next millennium: The continuous quality improvement approach. *Applied and Preventive Psychology, 8*, 247-255.
- Borduin, C. M. (1999). Multisystemic treatment of criminality and violence in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 242-249.

- *Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 35, 105-114.
- *Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63, 569-578.
- Borduin, C. M., & Schaeffer, C. M. (1998). Violent offending in adolescence: Epidemiology, correlates, outcomes, and treatment. In T. P. Gullotta, G. R. Adams, & R. Montemayer (Eds.), *Delinquent violent youth: Theory and interventions* (pp. 144-174). Newbury Park, CA: Sage.
- Borduin, C. M., Schaeffer, C. M., & Ronis, S. T. (2003). Multisystemic treatment of serious antisocial behavior in adolescents. In C. A. Essau (Ed.), *Conduct and oppositional defiant disorders: Epidemiology, risk factors, and treatment* (pp. 299-318). Mahwah, NJ: Erlbaum.
- Borenstein, M., Rothstein, H., & Cohen, J. (1997). *Power precision*. Hillsdale, NJ: Erlbaum.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Brown, T. L., Borduin, C. M., & Henggeler, S. W. (2001). Treating juvenile offenders in community settings. In J. B. Ashford, B. D. Sales, & W. H. Reid (Eds.), *Treating adult and juvenile offenders with special needs* (pp. 445-464). Washington, DC: American Psychological Association.
- *Brown, T. L., Henggeler, S. W., Schoenwald, S. K., Brondino, M. J., & Pickrel, S. G. (1999). Multisystemic treatment of substance abusing and dependent juvenile delinquents: Effects on school attendance at posttreatment and 6-month followup. *Children's Services: Social Policy, Research and Practice*, 2, 81-93.
- *Brunk, M., Henggeler, S. W., & Whelan, J. P. (1987). Comparison of multisystemic therapy and parent training in the brief treatment of child abuse and neglect. *Journal of Consulting and Clinical Psychology*, 55, 171-178.
- Burns, B. J., Hoagwood, K., & Mrazek, P. J. (1999). Effective treatment for mental disorders in children and adolescents. *Clinical Child and Family Psychology Review*, 2, 199-254.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). London: Sage.
- Elliott, D. S. (Series Ed.). (1998). *Blueprints for violence prevention*. Boulder, CO: University of Colorado, Center for the Study and Prevention of Violence, Blueprint Publications.
- Farrington, D. P., & Welsh, B. C. (1999). Delinquency prevention using family-based interventions. *Children and Society*, 13, 287-303.
- Fergusson, D., Horwood, J., & Lynskey, M. (1997). Children and adolescents. In P. M. Ellis & S. C. D. Collings (Eds.), *Mental health in New Zealand from a public health perspective*. Wellington, New Zealand: Ministry of Health.
- Hedges, L. V. (1991). Statistical considerations. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 29-40). New York: Russell Sage.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents*. Pacific Grove, CA: Brooks/Cole.
- Henggeler, S. W., & Borduin, C. M. (1992). *Multisystemic therapy adherence scales*. Unpublished instrument, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina.
- *Henggeler, S. W., Borduin, C. M., Melton, G. B., Mann, B. J., Smith, L. A., Hall, J. A., Cone, L., & Fucci, B. R. (1991). Effects of multisystemic therapy on drug use and abuse in serious juvenile offenders: A progress report from two outcome studies. *Family Dynamics of Addiction Quarterly*, 1, 40-51.
- *Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 868-874.
- *Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, 60, 953-961.
- Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. (1993). Family preservation using multisystemic therapy: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies*, 2, 283-293.
- *Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance-abusing and dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research*, 1, 171-184.
- *Henggeler, S. W., Rowland, M. D., Pickrel, S. G., Miller, S. L., Cunningham, P. B., Santos, A. B., Schoenwald, S. K., Randall, J., & Edwards, J. E. (1997). Investigating family-based alternatives to institution-based mental health services for youth: Lessons learned from the pilot study of a randomized field trial. *Journal of Clinical Child Psychology*, 26, 226-233.
- Henggeler, S. W., Rodick, J. D., Borduin, C. M., Hanson, C. L., Watson, S. M., & Urey, J. R. (1986). Multisystemic treatment of juvenile offenders: Effects on adolescent behavior and family interaction. *Developmental Psychology*, 22, 132-141.
- *Henggeler, S. W., Rowland, M. D., Randall, J., Ward, D. M., Pickrel, S. G., Cunningham, P. B., Miller, S. L., Edwards, J., Zealberg, J., Hand, L. D., & Santos, A. B. (1999). Home-based multisystemic therapy as an alternative to the hospitalization of youth in psychiatric crisis: Clinical outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1331-1339.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment of antisocial behavior in children and adolescents*. New York: Guilford Press.
- Henggeler, S. W., Schoenwald, S. K., Liao, J. G., Letourneau, E. J., & Edwards, D. L. (2002). Transporting efficacious treatments to field settings: The link between supervisory practices and therapist fidelity in MST programs. *Journal of Child and Adolescent Psychiatry*, 31, 155-167.
- Hoffman, L. (1981). *Foundations of family therapy*. New York: Basic Books.
- Hollingshead, A. B. (1975). *The four-factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.
- Huey, S. J., Henggeler, S. W., Brondino, M. J., & Pickrel, S. G. (2000). Mechanisms of change in multisystemic therapy: Reducing delinquent behavior through therapist adherence and im-

- proved family and peer functioning. *Journal of Consulting and Clinical Psychology*, 68, 451-467.
- Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin*, 102, 187-203.
- Kazdin, A. E. (2000). Treatments for aggressive and antisocial children. *Child and Adolescent Psychiatric Clinics of North America*, 9, 841-858.
- Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, 66, 19-36.
- Lahey, B. B., Moffitt, T. E., & Caspi, A. (Eds.). (2003). *Causes of conduct disorder and juvenile delinquency*. New York: Guilford Press.
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic enquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mostellor (Eds.), *Meta-analysis for explanation: A casebook* (pp. 83-127). New York: Russell Sage.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Loeber, R., Farrington, D. P., & Waschbusch, D. A. (1998). Serious and violent juvenile offenders. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 13-29). Thousand Oaks, CA: Sage.
- Mann, B. J., Borduin, C. M., Henggeler, S. W., & Blaske, D. M. (1990). An investigation of systemic conceptualizations of parent-child coalitions and symptom change. *Journal of Consulting and Clinical Psychology*, 58, 336-344.
- Minuchin, P. P. (1985). Families and individual development: Provocations from the field of family therapy. *Child Development*, 56, 289-302.
- Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice. (1999). *Delinquency cases in juvenile courts, 1999*. Washington, DC: Author.
- Rutter, M., Giller, H., & Hagell, A. (1998). *Antisocial behavior by young people*. Cambridge, MA: Cambridge University Press.
- Schaeffer, C. M., & Borduin, C. M. (2003). *Moderators and mediators of therapeutic change in multisystemic treatment of serious juvenile offenders*. Unpublished manuscript.
- Scherer, D. G., Brondino, M. J., Henggeler, S. W., Melton, G. B., & Hanley, J. H. (1994). Multisystemic family preservation therapy: Preliminary findings from a study of rural and minority serious adolescent offenders. *Journal of Emotional and Behavioral Disorders*, 2, 198-206.
- Schoenwald, S. K., Henggeler, S. W., Brondino, M. J., & Rowland, M. D. (2000). Multisystemic therapy: Monitoring treatment fidelity. *Family Process*, 39, 83-103.
- Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services*, 52, 1190-1197.
- Schoenwald, S. K., Ward, D. M., Henggeler, S. W., Pickrel, S. G., & Patel, H. (1996). Multisystemic therapy treatment of substance abusing and dependent adolescent offenders: Costs of reducing incarceration, inpatient, and residential placement. *Journal of Child and Family Studies*, 5, 431-444.
- **Schoenwald, S. K., Ward, D. M., Henggeler, S. W., & Rowland, M. (2000). MST versus hospitalization for crisis stabilization of youth: Placement outcomes 4 months post-referral. *Mental Health Services Research*, 2, 3-12.
- Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261-281). New York: Russell Sage.
- Szapocznik, J., Rio, A., Murray, E., Cohen, R., Scopetta, M., Rivas-Vasquez, A., Hervis, O., Posada, V., & Kurtines, W. (1989). Structural family versus psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology*, 57, 571-578.
- U.S. Public Health Service. (2001). *Youth violence: A report of the Surgeon General*. Washington, DC: Author.

Received March 4, 2003

Revision received January 2, 2004

Accepted February 14, 2004 ■